Predicting Vessel Tracks in Waterways Using Bi-LSTM and Transformer Models

Finn Minßen
Institute for the Protection of Maritime Infrastructures
German Aerospace Center (DLR)
Bremerhaven, Germany
finn-matthis.minssen@dlr.de

Matthias Steidel
Institute of Systems Engineering for Future Mobility
German Aerospace Center (DLR)
Oldenburg, Germany
matthias.steidel@dlr.de

Arto Niemi
Institute for the Protection of Maritime Infrastructures
German Aerospace Center (DLR)
Bremerhaven, Germany
arto.niemi@dlr.de

***Abstract*—This paper proposes an approach to predict vessel tracks in waterways using a bidirectional Long Short-Term Memory (Bi-LSTM) and a transformer model. For this purpose, the positions of the buoys along the Elbe and Weser rivers delimiting the waterway will be determined and merged with Automatic Information System (AIS) data. Additionally, tide data as well as weather information will be used to train the model. The created models are then trained with and without tide and weather data and evaluated against each other to check the influence of the additional attributes. Our results show that the lowest average error for each predicted feature is achieved with the Bi-LSTM, where no tide and weather data were used for training. Also the transformer model reaches lower prediction errors than a linear prediction, which is used as a baseline model.**

***Index Terms*—Track prediction; bidirectional LSTM; Transformer; AIS; Tide data; Weather information**

I. Introduction

The shipping industry is a important section of the global economy, with around 80% of global trade by volume carried out by sea [1]. This high volume leads to dense traffic, especially in coastal regions, due to the limited and narrow routes that can be navigated by vessels. Accidents such as that of the container ship Mumbai Maersk, which ran aground off Wangerooge, or the vessel that collided with a wind turbine in the north see, pose a major risks [2] [3]. To avoid these types of incidents, predicting vessel tracks for maritime situational awareness is becoming increasingly important, as these predicting can be used to detect anomalous vessel behavior [4].

In this paper, we introduce a bidirectional Long Short-Term Memory (Bi-LSTM) as well as a transformer model to predict vessel tracks in waterways. We use historical AIS data, tide information, weather data as well as topological information as inputs for these models. As will be described in section III-B, the topological information is used for representing the waterways as grids following the approach from Steidel et al. [5]. We then compare the results with a linear model consisting of a simple dense layer with a linear activation function. This model is used as a baseline because it makes simple and reasonable predictions.

These experiments are motivated by previous works. Steidel et al. [5] use kernel density estimation for predicting vessel tracks. However, their approach predict vessel behavior independently of the behavior in the previous corridor, which can lead to great variations for each of the considered parameters. Our approach predicts track through multiple corridors. Steidel et al. also considered enhancing the prediction performance by including weather information. As a recent review paper by Zhang et al. [6] point out, combining several data sources to predict vessel tracks is not yet sufficiently researched. We compare the prediction accuracies for models with and without the weather characteristics: wind speed, wind direction, wave height and tidal data. To the best of our knowledge, transformer models have so far only used AIS data for track prediction [4]. Therefore, combining additional data with AIS data to train a transformer model for track prediction is a novel experiment.

This paper is organized as follows. Section II gives a more thorough description of recent works in vessel track prediction. Our approaches are detailed in Section III. Part III-A introduces the Bi-LSTM and transformer models and Part III-B describes the input data sources. Section IV first describes how we prepared the input data and then provides the results. The results show that the average prediction errors obtained with the Bi-LSTM and the transformer models are lower than that of the linear model. This paper ends with section V where the presented results will be discussed.

II. Related Work

The prediction of vessel tracks has been intensively researched in recent years. Zhang et al. [6] summarize 57 different works covering this area. Their summary shows that machine learning has been becoming increasingly important since 2020. State-of-the-art models within research often use variations of LSTM models to predict the expected vessel movement.

One approach introduced by Mehri et al. [7] trains separate LSTM models for each type of vessel. To train these models, AIS data were used from November 2017 to the end of December 2017 from the eastern coast of the United States of America. In addition, geographic information was used to simplify vessel tracks. The models were then evaluated using the Root Mean Square Error (RMSE) and the point-wise horizontal error and compared with an ordinary LSTM model. The results show that the developed method receives a lower point-wise error in predicting the tracks accuracy with a range of up to 2 kilometers than an ordinary LSTM model.

Forti et al. [8] propose an encoder-decoder LSTM model that receives up to 20 previous positions to predict the next 20 vessel positions at two-minute intervals. To train the model, the authors used AIS data from June to September 2018 from the port of Piombino to Portoferraio on the island of Elba, Italy.

Also variations of LSTM models were used to predict ship track. Yang et al. [9] designed a Bi-LSTM to predict short-term vessel predictions. The model was trained with AIS data from 1364 ship tracks collected during one month in the Taiwan area. It was then evaluated by comparing the Mean Absolute Error, RMSE and the Mean-Absolute-Percentage Error with various other methods, such as simple LSTM and Recurrent Neural Network (RNN) models. Therefore, 10 vessel tracks were randomly compared. The results showed that the BI-LSTM has the lowest error in all categories compared to the other models. Based on this, Yang et al. [9] state that the developed model can accurately predict short-term trajectories.

A Bi-LSTM model can also be found in the work by Liu et al. [10]. They developed a series of routing algorithms where a Bi-LSTM was augmented with an attention mechanism to predict the next position of a vessel along the trajectory. The attention mechanism should make the prediction more accurate by better learning the dependencies of the AIS data. This method was trained with AIS data from fishery vessels along the east coast of China from May 2015 to May 2018. The results showed that the methods predict vessel positions to an error of less than 300 meters after one hour and an error of 2.73 kilometers after 9 hours when using an iterative process.

A different approach for trajectory predicting was developed by Nguyen et al. [4]. They state that standard deterministic approaches such as LSTMs cannot capture the multi-model patterns involved in AIS data and are therefore ineffective for trajectory prediction. With this assumption in mind, Nguyen et al. [4] propose a transformer structure to deal with the multi-model nature of AIS tracks. The transformer model contains 8 layers, each with 8 attention heads, and is being trained and tested with AIS data along the Danish coast during the first three months of 2019. The model was then evaluated, among others, against the sequence-to-sequence LSTM model proposed by Forti et al. [8]. Thereby, a significantly lower error in the forecast, measured with the haversine distance, is shown both in the first three hours and after 10 hours. Nguyen et al. [4] state that the developed model is more suitable to capture the multi-model nature of AIS-data and extract useful information in historical data than the compared models.

Much of existing research uses solely the vessel location information provided via the AIS to predict vessel tracks. Although, other factors such as weather conditions, tides and regional geographical characteristics also influence the track taken. As Zhang et al. [6] pointed out, combining several data sources to predict vessel tracks is not yet sufficiently researched.

III. MATERIALS & METHODS

A. Machine Learning Models

LSTM models are widely used for predicting vessel tracks. They are a type of recurrent neural network that utilize gated units to selectively control the flow of information within the network. LSTMs consist of a memory cell and three types of gates: input gate, forget gate, and output gate. The input gate regulates the flow of new information into the cell, the forget gate controls the flow of information out of the cell, and the output gate determines the information flow from the cell to the next hidden state. These gates allow selectively remembering or forgetting information, through which LSTMs are better able to model long-term dependencies in sequential data than regular RNNs. [11]

To extend the capabilities of LSTMs, Bi-LSTMs incorporate information from both past and future time steps of the input sequence. Thereby, the architecture consists of two LSTM layers, one processing the sequence in a forward direction and the other in a backward direction. The outputs of these layers are concatenated to produce the final output, allowing access to information from both temporal directions and improving prediction accuracy compared to regular LSTM models.

LSTMs process input sequences sequentially, updating their internal state one element at a time. This sequential nature limits parallelization and can lead to increased computational complexity and training time, especially for long sequences. To address this limitation, Vaswani et al. [12] propose the transformer model, which operates without any recurrence. The introduced transformer model employs a self-attention mechanism that enables efficient capturing of global dependencies by processing the entire input sequence in parallel. This attention mechanism calculates similarity scores for all pairs of positions in the sequence, allowing the model to learn deeper dependencies compared to LSTM models. The self-attention mechanism is combined with a feed-forward network, as well as normalization layers, and embedded in encoder and decoder layers. These components collectively enable the transformer model to effectively capture long-term dependencies which made it successful for time series predictions.

In this work, we want to compare both a Bi-LSTM and a transformer model to predict vessel tracks inside waterways. Therefore, we employed a Bi-LSTM model where each LSTM has 128 units followed by a dropout layer and a dense layer that predicts the future attributes. Overall, the model has at least 155,078 trainable parameters. This model will be compared with a transformer model that contains three transformer encoder layers, based on the architecture introduced by Vaswani et al. [12]. Using this architecture, the model contains at least 24,286 trainable parameters. This encoder layer is also followed by a dense layer predicting the future attributes. The linear model, to which the Bi-LSTM and the transformer model are compared, consists of a simple dense layer with a linear activation function.

B. Maritime Data

AIS data are based on the automatic exchange of data between vessels regarding their characteristics and positions with other vessels in their area. AIS data is collected, on the one hand, from AIS base stations that monitor traffic in specific areas and, on the other hand, from satellites that collect data on a global scale. Since the use of AIS is mandatory for specific vessels, these data can be used to obtain a view of shipping traffic all over the world and allows creating vessel tracks [13]. From the AIS data, we extract the Speed Over Ground (SOG), Course Over Ground (COG), vessel position (longitude, latitude), the vessel identification number (MMSI) and the time the message was sent.

Based on the approach of Steidel et al. [5], it is assumed that vessels in the area of interest have to sail within the waterways, because only there a minimum depth is ensured. Waterways are bounded by starboard and port buoys that structure traffic within the waterways. These buoys are placed at indeterminate intervals and at crossings to help navigate within the waterways. Within these waterways, vessels are obliged to sail as close as possible to the starboard buoy [14]. To extract the required information, we divide the waterways into a grid, where four buoys always form a cell, as shown in Fig 1. With this approach, a waterway can be viewed as a sequential series of cells and all AIS positions occurring within these cells can be filtered. This filtered AIS data is then used to create continuous AIS tracks along the considered waterways. An AIS track is defined as a series of AIS messages for a particular MMSI received within one minute of the previous message. If the interval between received messages is longer than one minute, a new track is created. From this AIS tracks transition points (TP) at which a vessel enters a cell, i.e., passes through the starboard and port buoys of the respective cell, can be calculated. Therefore the positions before and after entering the cell are interpolated linearly. Further, the distance to the starboard buoy (d_s) at the crossing can be calculated. This measurement simplifies the prediction, since instead of a position including latitude and longitude, only one value, the distance, has to be predicted.

The TP is also used to determine the angle (β) at which the vessel moves from one cell to the next. Therefore, TP, the position of the starboard buoy which the vessel crosses by entering the cell as well as the first position inside the new cell is used. These three points form a triangle, from which β can be calculated. At cells where waterways cross, β may suggest which waterway a vessel will follow.

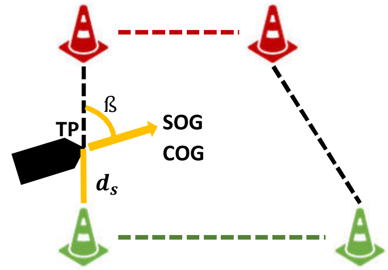

Fig. 1. Visualization of the extracted information from the waterway. The red cones show the port limits and the green ones the starboard limits of the waterway. TP indicates the transition point when the vessel enters the cell, while d_s indicates the distance to the starboard board buoy and β the angle at which the vessel enters the cell. Further SOG and COG are linearly interpolated for each TP.

Furthermore, it is assumed that tide information has an influence on the prediction accuracy. Therefore, tide information from data provided by the European Center for Medium-Range Weather Forecasts (ECMWF) is used [15]. These data consist of records of the water surface level from buoys located at specific positions along the coast and in rivers. The position of the recording buoys considered in this work are displayed in Fig. 2. Depending on the position of the track, the water surface level from the two recording buoys, which are referred to as b_1 and b_2, located at the beginning and end of the track, are added to the AIS track, by interpolating the water levels based on the time at which the vessel crosses the TP.

Fig. 2. Location and name of the six buoys that measure the water level in the area of interest. (Bremen, Wilhelmshaven, Wangerooge, Alte Weser, Cuxhaven and St Pauli)

Since the influence of weather data on the prediction will also be investigated, the characteristics: wind speed (w_s), wave height (w_h) and wind direction (w_d) will be extracted from the ERA5 dataset provided by Copernicus [16]. These are then attached to the TPs based on time and position.

In summary, a track is represented as a series of transition points TP, each containing the attributes: $\{d_s, SOG, COG, \beta, b_1, b_2, w_s, w_h, w_d\}$.

IV. Results

A. Data

In this paper, we use commercially available AIS data that were collected from terrestrial and satellite sources from

January 01, 2020 to April 30, 2020 along the German Bight and the Elbe and Weser rivers (Fig. 3).

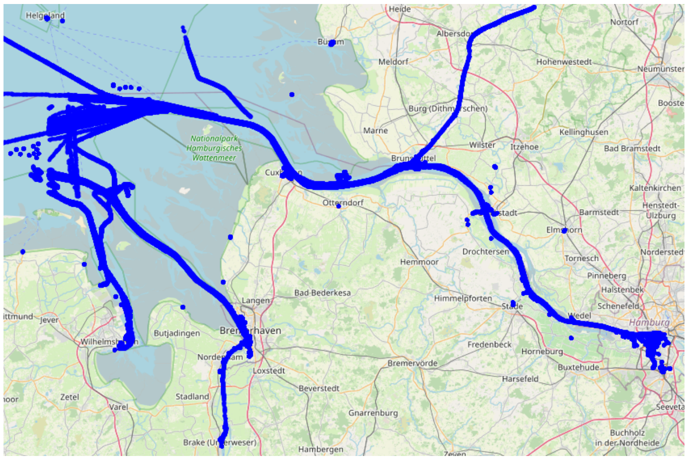

Fig. 3. AIS vessel positions in the area of interest

The AIS-data were then preprocessed by removing infeasible speed messages (2 knots $<$ SOG $<$ 30 knots). Furthermore, the process for generating TPs with the additional weather and tide characteristics as described in section III-B is summarized in Fig. 4

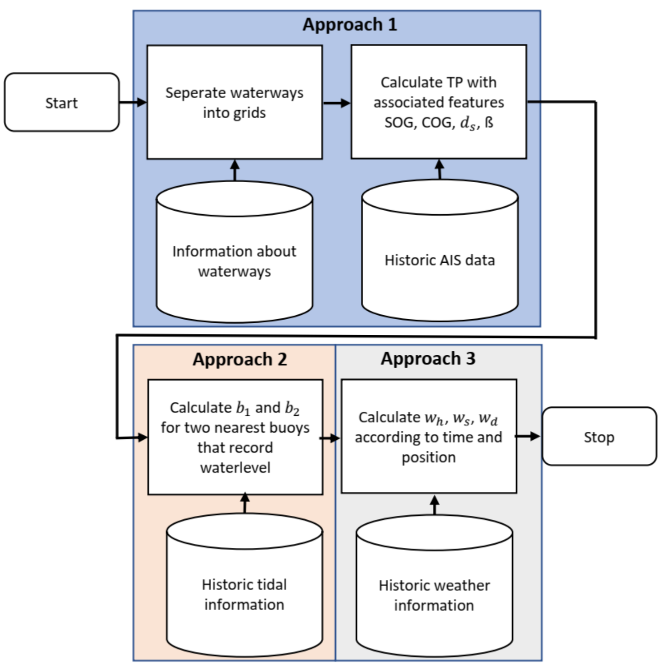

Fig. 4. Summary of the process for generating TPs in waterways with tide and weather characteristics. The data generated from the preceding approaches are incorporated into the succeeding approaches.

In total, 11,167 tracks with at least 20 and up to 60 transition points are used. Thereof, 2,877 tracks (26%) navigate along the Weser, 7,949 tracks (71%) along the Elbe and the remaining 341 tracks (3%) navigate to or from Wilhelmshaven. These data were then split into a training (80%) and a test (20%) data sets.

B. *Track prediction results*

All data-driven models were trained for 50 epochs using the Adam optimizer and the Mean Square Error (MSE) to measure the prediction error. In this process, the models are trained to predict the next TP based on the previous 10 TPs. The predicted TP is then used again with the previous nine TPs to predict the subsequent TP. Using this iterative approach, the subsequent five TPs were predicted. To evaluate the influence of the additional tide and weather data, each model was trained three times with different features. In the first approach, SOG, COG, d_s, and β were used exclusively to predict the vessel tracks. For the second approach, b_1 and b_2 were added to the existing features in the dataset, as well as the weather attributes for the third approach. Table I shows the average error for predicting the next five TPs compared to the ground truth.

TABLE I
AVERAGE PREDICTION ERROR OF THE SUBSEQUENT FIVE TPS COMPARED TO THE GROUND-TRUTH. d_s IS MEASURED IN METERS [M], SOG IS MEASURED IN KNOTS [KN] AND COG, β ARE MEASURES IN DEGREES [^{o}]. THE LINEAR MODEL WAS TRAINED FOR EACH FEATURE INDIVIDUALLY.

Input Data	*Model*	d_s *[m]*	*SOG [kn]*	*COG [^{o}]*	β *[^{o}]*
	Linear	178.09	0.74	9.97	18.93
Approach 1	Bi-LSTM	106.42	0.68	2.07	1.85
SOG, COG, d_s, β	Transformer	182.75	2.73	5.24	7.09
Approach 2	Bi-LSTM	197.46	1.02	2.16	2.03
b_1, b_2	Transformer	171.21	2.75	5.06	5.95
Approach 3	Bi-LSTM	153.97	1.25	2.11	2.18
w_h, w_s, w_d	Transformer	185.67	2.79	7.4	7.64

The average prediction error obtained with the Bi-LSTM model, which obtained the best result in comparison, can be divided according to the waterways along the Elbe and Weser rivers and Wilhelmshaven. This can be seen in Fig. II

TABLE II
AVERAGE PREDICTION ERROR DIVIDED BY WATERWAYS

Waterway	d_s *[m]*	*SOG [kn]*	*COG [^{o}]*	β *[^{o}]*
Elbe	100.57	0.65	2.17	1.87
Weser	115.15	0.71	1.90	1.76
Wilhelmshaven	105.15	0.82	2.22	2.64

As the results show, the lowest average prediction error for d_s is achieved along the Elbe river followed by the predictions along Wilhelmshaven and the Weser river. For SOG, the lowest error is again obtained along the Elbe. However, the difference between the best and the worst result applicable to Wilhelmshaven is only less than $0,2$ knots For the other features COG and β, the lowest error is received along the Weser River. The prediction results for these two features are, with less than one-degree difference, also close to each other. That the predictions to and from Wilhelmshaven is slightly worse than for the Elbe and Weser rivers, can be explained by the fact that the model was trained on only 3% of the total tracks in this waterway. Therefore, this waterway was underrepresented for the model. However, since the overall results of the predictions for the different waterways are close, it can be concluded that the presented concept for predicting ship tracks is generalizable.

V. Discussion

The results show that the Bi-LSTM trained in the first approach reaches the lowest average prediction error in all features measured. The lowest error using the transformer model is achieved in the second approach, where tide information (b_1, b_2) was added to the dataset. It is noticeable that the prediction results of the Bi-LSTM model become worse when tide and weather features were added. The transformer model, in contrast, decreased the prediction error by adding tide data. However, when adding weather data, the results also got worse.

It is noticeable that all models can produce erroneous data by predicting vessel tracks outside the waterway. In the respective best attempts, this occurs 0.7% of the time for the Bi-LSTM, 1.3% of the time for the transformer model, and 0.4% of the time for the liner model. This phenomenon occurs mainly in curve passages where the distance to the previous TP becomes significantly smaller as it can be seen in Fig. 5.

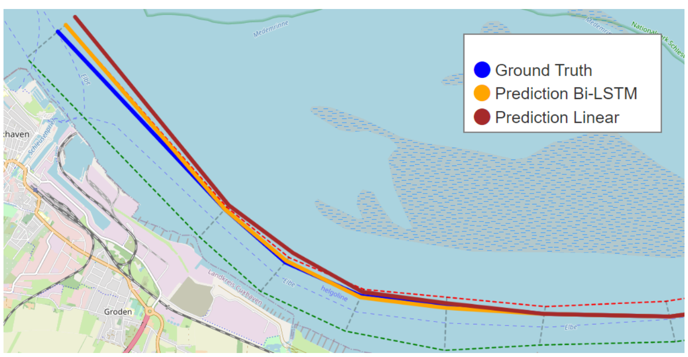

Fig. 5. Erroneous predictions outside the waterway.

Among the predictions, outliers with a large deviation from the ground truth were also predicted. For β and COG, the outliers mainly occur mainly for one location in the waterway along the Elbe river, which is displayed in Fig. 6. The waterway runs along Glückstadt and the buoys are arranged in such a way that it seems as if the vessel do not pass through the waterway from the north or south, but from the west or east. Since the arrangement of the buoys in the course of the waterway only give the impression at this intersection, the models cannot represent this transition well. Had the model actually predicted this outlier, this would have indicated that the model had overfitted and made predictions too close to the training data.

All in all, the results displayed in Table I show that the addition of tide information as well as weather characteristics does not automatically lead to an improvement of the predictions. Only the average prediction error for the transformer model decreases when tide data are added. Otherwise, adding

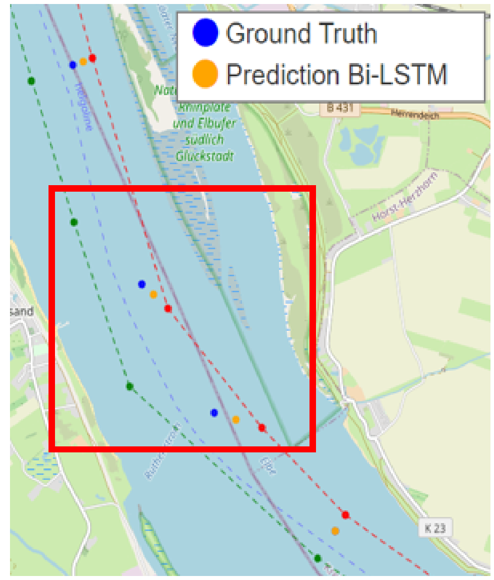

Fig. 6. Section of the waterway where outliers of COG and β are predicted due to the arrangement of the starboard and port buoys. Although the waterway leads north and south, the arrangement of the buoys indicates a more westerly and easterly course, respectively.

data lead to an increased average prediction error even though tide and weather data have a real impact on the navigation of vessels. One possible reason for that is that the new features may not be relevant or informative to the prediction task and therefore may introduce unwanted noise or bias to the model. That results in general can get worse with the addition of features has already been described by Kuhn et al. [17] and also by John et al. [18].

Overall, it should also be noted that the Bi-LSTM predicts the lowest average error. On the one hand, this result can be explained by the fact that it was also the most complex model and thus had more capacity to predict the dependencies. However, even the less complex transformer model makes reasonable predictions for the number of trainable parameters it possesses. It should be noted that approaches, to increase the complexity of the models to create a more comparable model lead to overfitting. Accordingly, the transformer architecture was left with the number of parameters. In particular, the strength of a transformer model is its ability to identify long-term dependencies [12]. In the presented approach, however, only the first following TP was predicted. Accordingly, the models did not have to learn large dependencies. For this kind of prediction, LSTM models are also well suited, as can be seen in the results.

In summary, both data-driven models show better results than linear prediction, except for SOG for the transformer model, which justifies the use of these models for this prediction task. However, the data-driven models could not use additional weather information to their advantage in order to make more accurate predictions. This issue may caused by the fact that including extraneous features hurts prediction performance of data driven models [17], [18]. This result adds nuance to the notion that including weather data would necessarily improve track predictions.

VI. Conclusion

In this paper, a concept for predicting Transition Points (TP) representing vessel tracks in waterways has been successfully developed, combining historical vessel position in the form of AIS data with weather information and tide data. Positions of buoys were extracted from sea chart information and were also combined with AIS data to create TPs. The data resulting from this concept was then used to train a Bi-LSTM and a transformer model. These models can be used to iteratively predict the subsequent transition points that represent the track a vessel will take. The model that predicts the most accurate vessel tracks was a Bi-LSTM model trained without tide and weather information, focusing only on AIS data combined with position of buoys that delimit waterways. However, also the transformer model predicted a lower average error than the linear prediction for d_s, COG and β.

In this paper, the transformer model overfitted as soon as more layers were added. Future work should experiment with different architectures, which may prevent this phenomenon. Solving it would improve predictions. In addition the developed concept can also be used for specific vessel types, which helps to assess the method for a real-world use. Further, the concept developed could also be used to detect anomalies. Thus, it could be elaborated to what extent the predictions can be used to detect anomalous vessel tracks in relation to the predicted features.

Acknowledgements

This paper uses data from Copernicus Climate Change Service information (accessed 2022). Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

References

[1] United Nations Conference on Trade and Development, *"Review of Maritime Transport 2021"*. "United Nations Publications", "2021".

[2] Z. Szymanska and M. Murray, “Maersk container ship runs aground off German island,” Reuters, News article, 2022. [Online]. Available: https://www.reuters.com/world/europe/container-ship-runs-aground-off-german-island-2022-02-03/

[3] M. Voytenko, “General cargo ship collided with wind turbine, damaged, north sea,” Maritime Bulletin, News article, 2023. [Online]. Available: https://www.maritimebulletin.net/2023/04/26/general-cargo-ship-collided-with-wind-turbine-damaged-north-sea/

[4] D. Nguyen and R. Fablet, “TrAISformer - A generative transformer for AIS trajectory prediction,” *CoRR*, vol. abs/2109.03958, 2021. [Online]. Available: https://arxiv.org/abs/2109.03958

[5] M. Steidel, J. Mentjes, and A. Hahn, “Context-sensitive prediction of vessel behavior,” *Journal of Marine Science and Engineering*, vol. 8, p. 987, 2020.

[6] X. Zhang, X. Fu, Z. Xiao, H. Xu, and Z. Qin, “Vessel trajectory prediction in maritime transportation: Current approaches and beyond,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 19 980–19 998, 2022.

[7] S. Mehri, A. A. Alesheikh, and A. Basiri, “A contextual hybrid model for vessel movement prediction,” *IEEE Access*, vol. 9, pp. 45 600–45 613, 2021.

[8] N. Forti, L. M. Millefiori, P. Braca, and P. Willett, “Prediction of vessel trajectories from AIS data via sequence-to-sequence recurrent neural networks,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 8936–8940.

[9] H. Li, H. Jiao, and Z. Yang, “AIS data-driven ship trajectory prediction modelling and analysis based on machine learning and deep learning methods,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 175, p. 103152, 2023.

[10] C. Liu, Y. Li, R. Jiang, Y. Du, Q. Lu, Z. Guo, and X. Zhang, “TPR-DTVN: A routing algorithm in delay tolerant vessel network based on long-term trajectory prediction,” *Wirel. Commun. Mob. Comput.*, vol. 2021, jan 2021. [Online]. Available: https://doi.org/10.1155/2021/6630265

[11] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, p. 1735–1780, nov 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735

[12] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30. Curran Associates, Inc., 2017.

[13] M. Stróżyna, W. Abramowicz, K. Węcel, D. Filipiak, and J. Małyszko, *Data Analysis in the Maritime Domain*. Poznań University of Economics and Business Press, 2022.

[14] IALA, “The IALA maritime buoyage system,” International Association of Lighthouse Authorities, Recommendation R1001, 2017.

[15] M. Sanne *et al.*, “Global sea level change time series from 1950 to 2050 derived from reanalysis and high resolution CMIP6 climate projections,” Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2022, accessed 2023. [Online]. Available: https://doi.org/10.24381/cds.a6d42d60

[16] H. Hersbach *et al.*, “ERA5 hourly data on single levels from 1979 to present.” Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2018, accessed 2023.

[17] M. Kuhn and K. Johnson, *Feature Engineering and Selection: A Practical Approach for Predictive Models*. Taylor & Francis Group, 2019, https://bookdown.org/max/FES/.

[18] G. H. John, R. Kohavi, and K. Pfleger, “Irrelevant features and the subset selection problem,” in *Machine Learning Proceedings 1994*, W. W. Cohen and H. Hirsh, Eds. San Francisco (CA): Morgan Kaufmann, 1994, pp. 121–129.